

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局



(10) 国际公布号
WO 2023/168998 A1

(43) 国际公布日
2023 年 9 月 14 日 (14.09.2023)

(51) 国际专利分类号:
G06V 20/40 (2022.01) *G06V 10/74* (2022.01)
G06N 3/04 (2023.01) *G06V 10/80* (2022.01)
G06N 3/08 (2023.01) *G06V 10/82* (2022.01)

圳市南山区高新区科技中一路腾讯大厦 35 层, Guangdong 518057 (CN)。

(72) 发明人: 郭卉 (GUO, Hui); 中国广东省深圳市南山区高新区科技中一路腾讯大厦 35 层, Guangdong 518057 (CN)。

(21) 国际申请号: PCT/CN2022/135034

(22) 国际申请日: 2022 年 11 月 29 日 (29.11.2022)

(74) 代理人: 华进联合专利商标代理有限公司 (ADVANCE CHINA IP LAW OFFICE); 中国广东省广州市天河区珠江东路 6 号 4501 房 (部位: 自编 01-03 和 08-12 单元) (仅限办公用途), Guangdong 510623 (CN)。

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
202210219364.3 2022 年 3 月 8 日 (08.03.2022) CN

(81) 指定国 (除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ,

(71) 申请人: 腾讯科技 (深圳) 有限公司 (TENCENT TECHNOLOGY (SHENZHEN) COMPANY LIMITED) [CN/CN]; 中国广东省深

(54) Title: VIDEO CLIP IDENTIFICATION METHOD AND APPARATUS, DEVICE, AND STORAGE MEDIUM

(54) 发明名称: 视频片段的识别方法、装置、设备以及存储介质

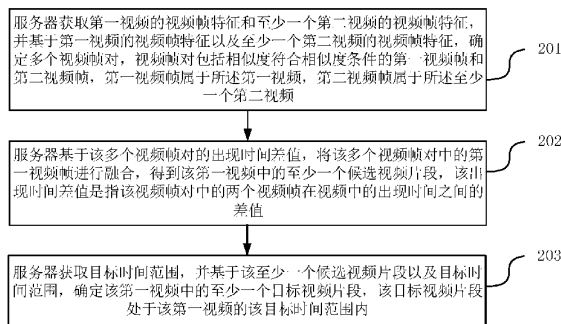


图 2

201 A server acquires a video frame feature of a first video and a video frame feature of at least one second video, and determines multiple video frame pairs on the basis of the video frame feature of the first video and the video frame feature of the at least one second video, the video frame pairs comprising first video frames and second video frames whose similarity meets a similarity condition, the first video frames belonging to the first video, and the second video frames belonging to the at least one second video

202 The server fuses the first video frames in the multiple video frame pairs on the basis of an occurrence time difference of the multiple video frame pairs, so as to obtain at least one candidate video clip in the first video, the occurrence time difference referring to a difference between the occurrence times of two video frames in the video frame pairs in the videos

203 The server acquires a target time range, and determines at least one target video clip in the first video on the basis of the at least one candidate video clip and the target time range, the target video clip being within the target time range of the first video

(57) Abstract: The present application discloses a video clip identification method and apparatus, a device, and a storage medium, which can be applied to scenarios such as video clip identification, artificial intelligence, and vehicle-mounted in a computer technology. By means of the technical solution provided by embodiments of the present application, the method comprises: acquiring a video frame feature of a first video and a video frame feature of at least one second video, and determining multiple video frame pairs on the basis of the video frame feature of the first video and the video frame feature of the at least one second video, the video frame pairs comprising first video frames and second video frames of which similarity meets a similarity condition, the first video frames belonging to the first video, and the second video frames belonging to the at least one second video (201); fusing the first video frames in the multiple

[见续页]

WO 2023/168998 A1

LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN,
MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE,
PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE,
SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ,
UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW。

(84) 指定国 (除另有指明, 要求每一种可提供的地区
保护): ARIPO (BW, CV, GH, GM, KE, LR, LS, MW,
MZ, NA, RW, SC, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧
亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT,
BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR,
HR, HU, IE, IS, IT, LT, LU, LV, MC, ME, MK, MT, NL,
NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ,
CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE,
SN, TD, TG)。

本国际公布:

— 包括国际检索报告 (条约第21条(3))。

video frame pairs on the basis of an occurrence time difference of the multiple video frame pairs, so as to obtain at least one candidate video clip in the first video, the occurrence time difference referring to a difference between the occurrence times of two video frames in the video frame pairs in the videos (202); and acquiring a target time range, and determining at least one target video clip in the first video on the basis of the at least one candidate video clip and the target time range, the target video clip being within the target time range of the first video (203).

(57) 摘要: 本申请公开了一种视频片段的识别方法、装置、设备以及存储介质, 可应用于计算机技术中的视频片段识别、人工智能以及车载等场景。通过本申请实施例提供的技术方案, 获取第一视频的视频帧特征和至少一个第二视频的视频帧特征, 并基于第一视频的视频帧特征以及至少一个第二视频的视频帧特征, 确定多个视频帧对, 视频帧对包括相似度符合相似度条件的第一视频帧和第二视频帧, 第一视频帧属于第一视频, 第二视频帧属于至少一个第二视频(201); 基于多个视频帧对的出现时间差值, 将多个视频帧对中的第一视频帧进行融合, 得到第一视频中的至少一个候选视频片段, 出现时间差值是指视频帧对中的两个视频帧在视频中的出现时间之间的差值(202); 获取目标时间范围, 并基于所述至少一个候选视频片段以及目标时间范围, 确定所述第一视频中的至少一个目标视频片段, 所述目标视频片段处于所述第一视频的所述目标时间范围内(203)。